DOCUMENT RESUME

ED 276 749 TM 860 700

AUTHOR Paulson, James A.
TITLE Latent Class Representation of Systematic Patterns in Test Responses.
INSTITUTION Portland State Univ., OR. Dept. of Psychology.
SPONS AGENCY Office of Naval Research, Arlington, Va. Personnel and Training Research Programs Office.
REPORT NO ONR-85-1
PUB DATE May 85
CONTRACT N00014-81-K-0564
NOTE 58p.
PUB TYPE Reports - Research/Technical (143)

EDRS PRICE MF01/PC03 Plus Postage.
DESCRIPTORS *Algorithms; Correlation; Estimation (Mathematics); Factor Analysis; Item Analysis; *Latent Trait Theory; Mathematical Models; Mathematics Tests; Maximum Likelihood Statistics; Psychometrics; Response Style (Tests); Scaling; *Test Construction
IDENTIFIERS *Latent Class Models; *Latent Structure Analysis

ABSTRACT

This paper discusses the use of latent class structure as a modelling framework for tests in which much of the data conforms to a relatively small number of systematic patterns. Application of this framework to the analysis of tests has been limited because available parameter estimation algorithms can only handle a relatively small number of items. This paper proposes a modification of the constrained maximum likelihood estimation algorithm of tests with many more items. A number of special cases of the latent class structure model and a series of problems which arise in the development of any latent class model are reviewed and discussed. A few problems with no satisfactory general solution are noted. The proposed procedures are used to develop and test models for signed-number addition test data gathered by Tatsuoka and Birenbaum. The models are shown to represent most features of the data well, including some significantly negative interitem correlations and differences between groups in covariance structure. The ability of these simple models to represent non-unidimensional data, and to represent differences between groups in covariance structure and means simultaneously, suggests that the latent class structure framework may be very useful in dealing with a number of troublesome psychometric issues. (Author/JAZ)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED276749

TM

LATENT CLASS REPRESENTATION OF SYSTEMATIC PATTERNS IN TEST RESPONSES

James A. Paulson

Psychology Department

Portland State University

May 1985

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

X This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

This research was sponsored by the Personnel and Training and Research Programs, Psychological Sciences Division, Office of Naval Research, under Contract No. N00014-81-K-0564, NR 150-466. I would like to thank Dr. Kikumi Tatsuoka for several helpful discussions of her signed-number arithmetic data and for permission to use the data.

BEST COPY AVAILABLE

TM 860 100

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER: Technical Report ONR85-1	2. GOVT ACCESSION NO.
3. RECIPIENT'S CATALOG NUMBER	
4. TITLE (and Subtitle): Latent Class Representation of Systematic Patterns in Test Responses	5. TYPE OF REPORT & PERIOD COVERED
6. PERFORMING ORG. REPORT NUMBER	
7. AUTHOR(s): James A. Paulson	8. CONTRACT OR GRANT NUMBER(s): N00014-81-K-0564
9. PERFORMING ORGANIZATION NAME AND ADDRESS: Psychology Department, Portland State University, Portland, OR 97207	10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS: NR 150-466
11. CONTROLLING OFFICE NAME AND ADDRESS: Personnel and Training Research Programs, Office of Naval Research, Arlington, VA 22217	12. REPORT DATE: May 1985
	13. NUMBER OF PAGES: 48
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office)	15. SECURITY CLASS. (of this report): Unclassified
	15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Latent class models	Systematic response patterns	Nonglobal maxima
Latent structure	Item equivalence	
Marginal maximum likelihood	Unidimensionality	
EM algorithm	Parameter invariance	
Signed-number arithmetic	Componential analysis	

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

This paper is concerned with the use of latent class structure as a modelling framework for tests in which much of the data conforms to a relatively small number of systematic patterns. Application of this framework to the analysis of tests has been limited because available parameter estimation algorithms can only handle a relatively small number of items. The paper proposes a modification of the constrained maximum likelihood estimation algorithm of Goodman (1974) which should extend the applicability of that algorithm to tests

20. ABSTRACT (cont'd.)
with many more items. A number of special cases of the latent class structure model are reviewed, including models which we propose for signed-number addition test responses. A series of problems which arise in the development of any latent class model are discussed and likelihood ratio tests for dealing with the problems are presented. A few problems with no satisfactory general solution are noted. The proposed procedures are used to develop and test models for signed-number addition test data gathered by Tatsuoka and Birenbaum (1979). The models are shown to represent most features of the data well, including some significantly negative interitem correlations and differences between groups in covariance structure. Negative interitem correlations imply that the data are not unidimensional. The ability of these simple models to represent non-unidimensional data, and to represent differences between groups in covariance structure and means simultaneously, suggests that the latent class structure framework may be very useful in dealing with a number of troublesome psychometric issues.

INTRODUCTION

The purpose of this paper is to present specific latent class structure models for testing mastery of signed-number addition. Representation of signed-number arithmetic test data requires a somewhat different set of constraints than other models employing latent class structure have used and some modifications of existing estimation procedures. These modifications may be of interest in their own right; they extend the applicability of latent class structure as a modelling framework, with little or no increase in the difficulty of the analysis.

The models for signed-number addition which I shall propose and the other models which will be considered in this paper all address the common problem of representing data in situations where it is believed that much of the data conforms to a relatively small number of systematic patterns. Probably the most well known situation of this kind is the Guttman scaling paradigm. In this situation items are assumed to be in a hierarchy of intensity or difficulty such that a person who endorses an item of given intensity or passes an item of given difficulty will endorse or pass all items of lesser intensity or difficulty on the scale. Other examples are suggested by cognitive analyses of test performance on subtraction problems by Brown and Burton (1978) and on signed number arithmetic problems by Tatsuoka and Birenbaum (1979). Both of these analyses of test

protocols found that many of the students who do not understand the concepts involved are not responding at random but are rather using incorrect rules which lead them to characteristic patterns of correct and incorrect responses. For example, if, in adding signed numbers, a student determines the signs correctly but always subtracts the smaller number from the larger in determining the absolute value of the sum, he or she will get all problems in which the addends have different signs correct, while missing all those whose addends have the same signs.

The latent class structure framework can be used to address two issues which arise in attempts to model situations like those just described. One issue concerns the problem of what to do about data that do not fit any of the hypothesized ideal patterns. In real applications there always is such data and, in the absence of a model capable of representing it, it must be dealt with by ad hoc means. In the context of Guttman scaling, Clogg and Sawyer (1981) argue persuasively that the ad hoc means often employed are unsatisfactory and that appropriate probabilistic generalizations of the ideal Guttman scale are much better.

The same remark applies to the cognitive analysis of systematic error patterns, where another issue arises. In order to deal with this kind of data it will often be necessary to go beyond unidimensional or hierarchical representations of the data. Consider the signed number example again. Another erroneous rule that often occurs is complementary to the one described above: some students always add the absolute values of addends to determine the absolute value of the sum. Students following

the latter rule miss exactly those items which students following the "always subtract" rule get correct and vice versa. It does not make sense to try to order the response patterns characterizing these two rules. Latent class analysis provides a convenient framework for representing this kind of nonhierarchical data.

Let us briefly review the elements of latent class models. The reader desiring a thorough introduction can consult Lazarsfeld and Henry (1968) for a clear, authoritative account of work up to 1968. The data to be accounted for are vectors of responses to items. In this paper we are only concerned with dichotomous item responses, although many of the ideas can be generalized to the case of polychotomous responses (cf. Goodman, 1974). It is assumed that every subject belongs to exactly one of a finite set of mutually exclusive and exhaustive latent classes. Theoretically, the distribution of the response vectors is to be accounted for by two sets of parameters and one key assumption. The two sets of parameters are the state probabilities,(ν_k), giving the distribution of subjects over the latent classes and the conditional probabilities of correct response to each item, given the respective states, (p_{kj}). The key assumption is that the responses are conditionally independent, given the state of subject. This implies that any relationships between items must be explained in terms of differences in the p_{kj}'s between classes. Models are specified by stipulating the number of classes and by placing constraints on the matrix of conditional probabilities.

For some time the problem of estimating parameters in latent class structures presented a real obstacle in the application of the latent class framework. It is necessary to employ iterative procedures in which

one selects a set of trial values, improves upon these values in the light of the data via some appropriate algorithm, and then repeats the process until (one hopes) the values stabilize at a good solution. McHugh (1956) derived the maximum likelihood estimators, but the solution applies only to the unconstrained model.

A great advance was achieved when Goodman (1974) described a particularly simple interative procedure which is guaranteed to converge to the maximum likelihood estimates under very general conditions. Goodman's procedure also has the virtue of automatically producing estimates of probabilities which fall in the unit interval; furthermore it is very easy to modify the procedure to satisfy a fair variety of constraints on the parameters. There is one problem which Goodman's procedure shares with McHugh's, however. Both procedures take as their data the frequency counts in the cells of the multi-way item-by-item ---by-item contingency table. For a relatively small number of items this presents no problem. But the number of cells grows exponentially with the number of items, so calculations which require dealing with all these cells become impractical very quickly as the number of items is increased. For example, data from a 20 item test will be discussed later in this paper; the contingency table for this test would have 2^{20} cells, i.e. more than a million. Fortunately, it is possible to formulate an algorithm which is logically equivalent to Goodman's, but which circumvents the problem of dealing with all possible cells in the n-way contingency table. This modified algorithm will be described in the next section of the paper. In the third section, the application of the algorithm to various constrained models will be discussed. These will

include models I am proposing for signed-number addition, which I will describe in that section. The fourth section of the paper will desribe the basic problems encountered in developing any latent class model and present statistical tests for dealing with these problems. The final section before the summary will then address these problems of model development with data on signed-number addition.

PARAMETER ESTIMATION

Let us introduce some notation and review formulas for some fundamental quantities.

x_{ij} = 1 if subject i is correct on item j; 0 if subject i is incorrect on item j;

n = the number of subjects;

J = the number of items;

ν_k = the probability of a subject being in class or state k;

s = the number of latent classes or states;

$\underline{x}_i$ = the vector of responses of subject i to items j = 1, .., J;

p_{kj} = the conditional probability that a subject in state k will respond correctly to item j;

$l_k(\underline{x}_i)$ = the conditional probability that a subject would give the response pattern $\underline{x}_i$, given that he or she is in state k; i.e. the conditional likelihood of response pattern $\underline{x}_i$, given state k;

$l^*(\underline{x}_i)$ = the weighted average of the conditional likelihoods over the states, called the marginal likelihood of $\underline{x}_i$;

ν_{ki} = the conditional probability that a subject is in state k, given that they make response pattern $\underline{x}_i$;

$L(\underline{x}_1,..,\underline{x}_n)$ = the overall marginal likelihood of all the response vectors $\underline{x}_1,...,\underline{x}_n$.

Basic Formulas

We need expressions in terms of the responses x_{ij} and the parameters, the ν_k's and p_{kj}'s, for the conditional and marginal likelihoods of the individual responses vectors, for the overall marginal likelihood, and conditional state probabilities, given a response pattern. These formulas are, respectively,

$$
\begin{aligned}
l_k(\underline{x}_i) &= \prod_{j=i}^{J} p_{kj}^{x_{ij}}(1-p_{kj})^{1-x_{ij}}, \\
l^*(\underline{x}_i) &= \sum_{k=1}^{S} \nu_k l_k(\underline{x}_i), \\
L(\underline{x}_1,\ldots,\underline{x}_n) &= \prod_{i=1}^{n} l^*(\underline{x}_i), \\
\nu_{ki} &= \frac{\nu_k l_k(\underline{x}_i)}{l^*(\underline{x}_i)} \quad .
\end{aligned}
\tag{1}
$$

The task of a marginal maximum likelihood estimation algorithm is to find the values of the parameters that maximize $L(\underline{x}_1,\ldots,\underline{x}_n)$. The EM algorithm of Dempster, Laird, and Rubin (1977) is a general approach to finding such algorithms. It can be used to justify the algorithms to be given now, or the logically equivalent algorithms of Goodman (1974).

Algorithm for the Unconstrained Model

The iterative process is particularly simple for the unconstrained model.

1. Choose a reasonable set of parameter values as your initial guess.
2. Compute the conditional state probabilities for each subject, given his or her response pattern, i.e., the v_{ki}'s.
3. Compute revised estimates of the state probabilities by the formula

$$v'_k = \frac{\sum_{i=1}^{n} v_{ki}}{n} \quad , \quad \text{for} \quad k=1,\ldots,s. \tag{2}$$

That is, the revised state probabilities are just the means of the conditional state probabilities, given the respective response patterns.

4. Compute revised estimates of the conditional probabilities of correct response to each item, given each state, by the formula

$$p'_{kj} = \frac{\sum_{i=1}^{n} v_{ki} x_{ij}}{\sum_{i=1}^{n} v_{ki}} \quad , \quad \text{for} \quad \begin{matrix} j=i,\ldots,J \\ k=1,\ldots,s. \end{matrix} \tag{3}$$

This formula might be thought of as dividing the estimated number of correct responses to item j given by subjects in state k, $\sum_{i=1}^{n} v_{ki} x_{ij}$, by the estimated number of subjects in state k, $\sum_{i=1}^{n} v_{ki}$.

5. Repeat steps 1 to 4 using the revised estimates of the ν_k's and the p_{kj}'s in step 1, until the estimates no longer change significantly.

Note that Equations 2 and 3 necessarily yield values between 0 and 1. This is an important feature of this algorithm which is not shared by other algorithms one might use, such as the Newton-Raphson method. Also note that the summations in Equations 2 and 3 are over subjects. These sums correspond to equivalent weighted summations over all cells in a J-way contingency table in the formulation of Goodman (1974). This is really the only difference between my formulation and Goodman's, but it is an important one from a practical point of view. It might be noted that in a similar context Bock and Aitken (1981) made the same sort of modification to an algorithm of Bock and Lieberman (1970) and found that it dramatically increased the number of items that could be handled effectively.

Algorithms for Constrained Models

It is easy to implement constraints that simply specify the values that some of the p_{kj}'s must have. One specifies these values as the initial values and then omits the revision step for the specified parameters at each iteration, so they never change. All other kinds of constraints can be implemented by carrying out all the steps of the unconstrained algorithm at each iteration, but adding a step at each iteration to deal with the constraint. Only constraints involving the p_{kj}'s will be considered here.

Suppose Θ is a parameter affecting several of the p_{kj}'s. The equation which results from setting the derivative of the logarithm of the marginal likelihood function equal to 0 can be rearranged to read

$$n \sum_j \sum_k \nu'_k \frac{(p'_{kj}-p_{kj})}{p_{kj}(1-p_{kj})} \cdot \frac{\partial p_{kj}}{\partial \Theta} = 0 \ . \tag{4}$$

While in general the presence of the $p_{kj}(1-p_{kj})$ factors in the denominators of each term makes equations like Equation 4 difficult to solve, there are enough special cases in which they simplify to cover a number of kinds of constraints one would like to be able to impose.

Suppose we would like to require that $p_{kj} = \Theta$ for one set of p_{kj}'s and $p_{kj} = 1-\Theta$ for another set, and that no p_{kj} outside of these sets depends on Θ. Then $p_{kj}(1-p_{kj})$ equals the constant $\Theta(1-\Theta)$ for all j,k such that the partial derivative $\partial p_{kj}/\partial\Theta$ is unequal to 0. The partial derivative is 1 for p_{kj}'s equal to Θ and -1 for those equal to $1-\Theta$. Let I_Θ be the set of indices j,k for which p_{kj} equals Θ and $I_{\bar{\Theta}}$ the set for which p_{kj} is $1-\Theta$. Then Equation 4 reduces to a linear equation in Θ whose solution is

$$\Theta = \frac{\sum_{j,k\in I_\Theta} \nu'_k p'_{kj} + \sum_{j,k\in I_{\bar{\Theta}}} \nu'_k (1-p'_{kj})}{\sum_{j,k\in I_\Theta} \nu'_k + \sum_{j,k\in I_{\bar{\Theta}}} \nu'_k} \tag{5}$$

That is to say, Θ is a plausible weighted average of the revised p'_{kj}'s which were computed with Equation 3 in the unconstrained phase of the iteration, with the weights being the revised state probabilities computed with Equation 2, the ν'_k's. To satisfy the constraints we replace the p'_{ij}'s involved with Θ or $1-\Theta$, as required, using the value of Θ obtained in Equation 5, and proceed to the next iteration.

In the case of simple equality constraints where we require that a set of p_{kj}'s be equal, the index set $I_{\bar{\Theta}}$ is empty and Equation 5 reduces to

$$\Theta = \frac{\sum_{j,k\varepsilon I_\Theta} \nu'_k \, p'_{kj}}{\sum_{j,k\varepsilon I_\Theta} \nu'_k} \tag{6}$$

If the p_{kj}'s involved in a constraint concern n_Θ items and just one state, the weight cancels and Equation 5 reduces to

$$\Theta = \frac{\sum_{j\varepsilon I_\Theta} p'_{kj}}{n_\Theta} \tag{7}$$

or

$$\Theta = \frac{\sum_{j\varepsilon I_\Theta} p'_{kj} + \sum_{j\varepsilon I_{\bar{\Theta}}} (1-p'_{kj})}{n_\Theta} \tag{8}$$

depending on whether the constraint is an equality constraint or a combination of equality and complementarity constraints.

Equality, complementarity, and fixed value constraints are useful as far as they go. However, they imply that each p_{kj} is influenced by only one parameter. This rules out models which characterize each p_{kj} in terms of conjoint effects of item and state parameters. Let $L(\chi)$ denote the logistic function, $(1 + e^{-\chi})^{-1}$. Let β_j be an item parameter affecting all p_{kj}'s involving item j and Θ_k a state parameter affecting all p_{kj}'s in state k. Let us constrain the p_{kj}'s to be the linear-logistic functions of the β_j's and Θ_k's:

$$p_{kj}(\beta_j, \Theta_k) = L(\beta_j + h_{kj}\Theta_k), \tag{9}$$

where the h_{kj}'s are constants specified in advance. Thus, the β_j's allow for overall differences between items which apply across states, whereas the Θ_k's may boost, diminish or not affect the conditional probability of correct response to an item, given state k, depending on how h_{kj} is specified for the item on the model.

The partial derivatives of the p_{kj}'s with respect to the β_j's and Θ_k's are

$$\frac{\partial p_{kj}}{\partial \beta_j} = p_{kj}(1-p_{kj}) \ , \tag{10}$$

$$\frac{\partial p_{kj}}{\partial \Theta_k} = p_{kj}(1-p_{kj})h_{kj} \ .$$

Therefore, Equation 4 simplifies somewhat, yielding the system of equations

$$\frac{\partial}{\partial \beta_j} \cdot \sum_{k=1}^{s} \nu'_k (p'_{kj} - p_{kj}) = 0 \ , \quad \text{for} \quad j=1,\ldots,J;$$

$$\frac{\partial}{\partial \Theta_k} \cdot \sum_{j=1}^{J} \nu'_k (p'_{kj} - p_{kj}) h_{kj} = 0 \ , \quad \text{for} \quad k=1\ldots,s. \tag{11}$$

Since the p_{kj}'s are nonlinear functions of the β_j's and Θ_k's, the system in Equation 11 is still somewhat cumbersome to solve exactly. However, if we replace the p_{kj}'s and the p'_{kj}'s by their logits, i.e., by

$$\gamma_{kj} = \log \frac{p_{kj}}{1-p_{kj}} = \beta_j + h_{kj}\Theta_k \ ,$$

we obtain a linear system of equations which ought to have approximately the same solution as Equation 11. We have not yet completed programming the algorithm incorporating these linear-logistic constraints, but models employing them may be attractive alternatives to existing models.

APPLICATIONS

The Latent Distance Model.

Most applications of constrained latent class models to date are special cases or variants of a model called the latent distance model by Lazarsfeld and Henry (1968). This model places the following constraints on the general model. First, the p_{kj}'s for a given item can have at most two values, a_j and b_j, where $0 \leq a_j \leq b_j \leq 1$. Second, the items can be ordered such that if the conditional probability of correct response in a given state takes on the lower value a_j for an item, it must take on the lower values for all the succeeding items. When all the a_j's are zero and all the b_j's are one, we have a pure Guttman scale. In various contexts, it may be desirable to place different constraints on a latent distance model, such as requiring all b_j's to be equal, or all b_j's to equal $1-a_j$. Clogg and Sawyer (1981) have reviewed applications of this sort in the scaling context.

Dayton and Macready (1976) have considered the application of the latent distance model to the analysis of hierarchies in the learning of different related concepts. They also consider models in which the ordering restriction on the items is relaxed. They developed their own maximum likelihood estimation algorithms for this relaxed version of the latent distance model.

The Quasi-Independence Model

Goodman (1975) has proposed an alternative way to introduce a probabilistic element into situations characterized by prominant

systematic patterns of response. This model, the quasi-independence model, constrains the conditional probabilities of correct response, given a systematic state, to be zero or one. One special state is added to account for unsystematic patterns; the conditional probabilities in this state are left unconstrained. Thus, for any given item there are three values which the p_{kj}'s can take: 0, 1, and the p_{kj} for the unsystematic state. Hence, this model does not satisfy the constraints of the latent distance model. In the opinion of Clogg and Sawyer (1981), the quasi-independence model may be preferable to the latent distance model for analyzing scalability of response patterns.

A Model for Signed-Number Addition

Dayton and Macready (1980) have generalized the quasi-independence model to allow response errors in the systematic states. The model I am going to propose now for signed-number addition test performance is very similar to this latter model of Dayton and Macready's. The important way it differs from that model is that it does not assume that the systematic states are ordered. It is not a scaling model.

In order to correctly add signed-numbers, one must learn two things: first, the sign of the larger number determines the sign of the sum; second, the absolute value of the sum is the sum of the absolute values if the signs of the addends agree but is the difference between the larger and the smaller absolute values if they disagree. In analyzing signed-number arithmetic test data, Tatsuoka and Birenbaum (1979) found that one gets a much clearer idea of what is going on if one scores the sign component of the response and absolute value component of the response separately. When one does so, clear patterns of error emerge in

the data on both components, for many of the subjects. The model I am proposing consists of a separate latent class model for each component score.

The model for each component postulates a mastery state, a biased random guessing state in which the probability of guessing correctly varies with the type of item, plus some states corresponding to systematic misconceptions. Three things can lead a subject to respond incorrectly in this kind of model: a subject in the random guessing state can guess incorrectly; a subject can harbor a misconception which systematically leads him or her to respond incorrectly to some types of problems and correctly to others; finally, subjects in the mastery state, or with a misconception that would ordinarily lead to a correct response to a given item, can make an error in the execution of their usual rule.

The parameters of the model for each component are the proportions of subjects falling in each state (ν_k's), the probabilities of execution errors in each state (ε_k's), and the probabilities of correct guesses to the respective item types in the random guessing state (p_{rj}'s). Table 1 summarizes specific models for the two components. It specifies the conditional probability of correct response for each type of addition problem included in data collected by Tatsuoka and Birenbaum (1979) from junior high school students who had just begun to study signed-numbers. It should be noted that there were four items of each type, so there were twenty items in all.

INSERT TABLE 1 ABOUT HERE

The models presented in Table 1 embody various constraints. All the item parameters in the Mastery State are assumed to be equal in the models for both components. In the states corresponding to systematic error patterns, a complementarity constraint is imposed: each item parameter in the state is either $1-\varepsilon$ or ε, depending on whether subjects in the state tend to get the item right or wrong. In the random guessing states, items of the same type are all assumed to have the same probability of correct response. More general models in which these constraints are relaxed will also be considered below in discussing the process of model development.

MODEL DEVELOPMENT

It will be instructive to consider the process of model development in some detail. The development of the models for the two components of responses to signed-number addition test items will be used to illustrate the process. But first it is desirable to review the statistical tests and an idealized sequence of steps one might go through in the development of any latent class model.

The likelihood ratio test

Let L_1 represent the maximum value of likelihood function described in Equation 1 for a model with $\underline{s}$ free parameters. In other words, L_1 is the value of the likelihood function evaluated with the parameter values set equal to their marginal maximum likelihood estimates. Let L_2 represent the corresponding maximum value for a more general model which subsumes the model used to obtain L_1 as a

special case. Suppose the more general model has q free parameters, where q is greater than s. The likelihood ratio criterion is the ratio

$$\Lambda = \frac{L_1}{L_2} \quad . \tag{12}$$

Under the null hypothesis that the more restricted model holds, the statistic

$$-2\log\Lambda = 2(\log L_2 - \log L_1) \tag{13}$$

asymptotically has a chi-squared distribution with $q - s$ degrees of freedom. By suitable specification of subsuming and subsumed models, one can address most of the basic problems in model development we shall face.

Some basic problems

The first question which arises in developing a latent class model for a set of data is how many latent classes or states the model should have. Once it has been decided how many states there should be, questions arise concerning the detailed specification of a model for the conditional probabilities of correct response in each state. In many applications, such as the signed-number addition test to be discussed in detail below, the test items have been constructed in a way that suggests a natural partition of the items into equivalence classes. In such cases it is desirable to test the appropriateness of this characterization of items before proceeding to a more detailed specification of the model. In fact,

models which require only that hypothetically equivalent items have the same parameters within a given state may be better than the completely unconstrained models for determining the number of states to employ. If the outcomes of the preliminary states of model development are favorable, there are various more specific questions to be addressed. For example, it is at this stage one would test the fit of the specific models presented in Table 1 for the components of responses to signed-number addition problems. If one obtains a reasonable fit for such a model to data from a given group, further questions arise. Are the conditional probabilities one obtains invariant from group to group, so that group differences can be attributed to differences between the groups in distribution over states? In the same group, are the conditional probabilities invariant over time, so that changes in performance on the test can be interpreted as transitions between states? All of the questions mentioned so far can be addressed using the likelihood ratio test. Perhaps it would be good to say a few words about each specific kind of test.

Determining the number of states. In order to use Equation 13 in any particular problem, it is necessary to specify the restricted model which serves as the null hypothesis and the more general model to be taken as the alternative hypothesis. In most applications regarding frequency distributions in contingency tables, the alternative model is implicitly taken to be the observed distribution of frequencies in the cells of the table. This choice is not feasible in the analysis of the latent class structure of a test unless the number of items is very small. The number of cells in the multiway contingency table for a test with k items is

2^k. Thus, for a test with ten items there are 1,024 cells in the contingency table. The sample size required to make the chi-squared distribution a reasonable approximation to the distribution of the test statistic might be tens or hundreds of thousands. For a test with twenty items there are more than a million cells, so the choice of the observed distribution as the alternative is infeasible for any conceivable sample size.

The first step in determining the number of states is to see if at least two states are required. Ordinarily, it would be safe to assume they are required and skip this step, because the null hypothesis that one state is sufficient is equivalent to the hypothesis that the item responses are all mutually independent. However, one can formally test it by computing $2(\log L_2 - \log L_1)$, where L_2 is the maximum of the likelihood function under the hypothesis of an unconstrained two-state model and L_1 the corresponding value for the one-state model. The degrees of freedom for the chi-squared test equals the number of items, plus one. That is one degree of freedom for each item parameter in the second state, plus the parameter needed to specify the distribution of subjects over the two states.

In general, if it has been determined that at least k states are needed, one computes $2(\log L_{k+1} - \log L_k)$, where L_{k+1} is the maximum of the likelihood function under the hypothesis that $k+1$ states are needed and L_k is the maximum under the null hypothesis that k states suffice. Under the null hypothesis, this statistic also is asymptotically distributed as chi-squared with degrees of freedom equal to the number of items plus one. Ideally, this testing process would continue until one

would not need to reject the null hypothesis. However, the number of parameters in the unconstrained model grows directly with the number of states. If there are s states and J items, then the number of free parameters is $Js + s-1$. The number of parameters that can be estimated in principle cannot be larger than the number of distinct response patterns, which in turn cannot exceed the number of subjects, and is usually much smaller than that. Thus, there is a definite limit to how far one can go with completely unconstrained models in determining the number of states needed to represent the data.

Testing the equivalence of sets of items. Let us say that a set of items are equivalent, if, for every state, the conditional probability of correct response, given the state, is the same for all the items in the set. Suppose that it has been determined by the process described in the last paragraph that k states with unconstrained item parameters are sufficient to represent the data. Suppose that the J items on the test consist of m sets of items which are hypothesized to be equivalent within sets. The hypothesis of equivalence can be tested using the likelihood ratio chi-squared statistic in Equation 13, with L_1 being maximum of the likelihood function calculated under the constraint that item parameters within equivalence classes are equal, and L_2 the unconstrained maximum for a model with k states. The constrained model only requires m parameters per state, whereas the unconstrained model has J parameters per state. Therefore, the degrees of freedom for the chi-squared statistic is $(J-k)s$. Note that the parameter values needed to compute the constrained maximum L_1 are easily obtained by the modification to the unconstrained algorithm given in Equation 7.

Testing a specific parametric model. In order to test specific models like the ones proposed above for signed-number addition items, we let L_1 be the maximum of the likelihood for the model with the specific constraints on the item parameters and L_2 the maximum for a more general model which subsumes the specific model as a special case - the usual pattern for setting up the likelihood ratio test. In situations where the items fall into at least nominally equivalent subsets, the investigator has some choice concerning the general model. Presumably, the issue of the number of states has been settled by the methods outlined in previous paragraphs. The number of states ought to be the same as in the specific model. Whether one ought to use the completely unconstrained model or the model constraining item parameters within nominally equivalent subsets to be equal is another question.

Let L_2 be the maximum for the model with the equivalent items constraint and L_3 the maximum for the unconstrained model. The identity

$$2(\log L_3 - \log L_1) = 2(\log L_3 - \log L_2) + 2(\log L_2 - \log L_1) \qquad (14)$$

shows that the likelihood ratio chi-squared statistic for testing the fit of the specific model relative to the unconstrained model can be broken into two components: the statistic for testing the equivalence of items hypothesis, and the statistic for testing the fit of the specific model relative to the equivalent items model. Since the statistics for the two components of the goodness-of-fit statistic provide information about different questions, it would be a good idea to use both of them.

Testing the invariance of item parameters. There are several ways one might approach the problem of seeing if the entire matrix of item parameters (i.e., conditional probabilities of correct response, given state) varies significantly from group to group, or from one occasion to another. One way is a single-sample approach which regards one set of parameters as given population values and the other set as estimates based on a single sample. Other ways involve a two-sample approach which regards both sets of parameters as sample estimates. Two-sample comparisons involving data obtained from the same subjects on different occasions require a different test than comparisons involving independent groups of subjects.

To apply the likelihood ratio test in the single-sample method, let the item parameter estimates in the first group be regarded as the population values. Let L_1 be the maximum of the likelihood function calculated for the second group under the constraint that the item parameters are identical to those obtained from the first group; only the state probabilities are allowed to vary in determining L_1. Let L_2 be the maximum obtained in the second group in a complete recalibration of the model in which both the state distribution and item parameters are estimated. For purposes of comparability, the item parameter estimates should be subject to the same constraints used to obtain corresponding estimates in the first group. Equation 13 may then be used to compute the chi-squared statistic, which will have degrees of freedom equal to the number of free item parameters estimated in the calibration used to obtain L_2.

The two-sample approach with independent groups is carried out as follows. Let L_1 be the maximum of the likelihood function for the combined groups. Let L_a and L_b be the maxima obtained in separate calibrations in groups a and b under the same constraints used to obtain L_1. Then the joint maximum likelihood for the combined groups, allowing their parameters to differ, is given by

$$L_2 = L_a L_b.$$

The statistic in Equation 13 becomes $2(\log L_a + \log L_b - \log L_1)$. It has K degrees of freedom, where K is the number of free parameters in each separate maximization. Clearly, K depends on the nature of the constraints in the model. The extension of this idea to the comparison of several groups is straightforward.

The single-sample and two-sample approaches have different advantages and disadvantages, which would lead one to use the first in some circumstances and the second in others. The single-sample approach does not require the raw data from the first group, just the item parameters. It might therefore be applicable in situations where it would be difficult or impossible to use the two-sample method. The two-sample approach takes sampling variability in both sets of parameters into account, so it would be preferred in most situations where all the data from both samples is available.

In order to assess the invariance of item parameters in repeated testing of the same subjects on two different occasions we can proceed as follows. For purposes of analysis, we regard the performance of subjects on the two testings as one long test, with the first half of the "items" being the first test and the second half of the "items" the second test. Since subjects may change their state between the test administrations, it is necessary to expand the states to include all possible pairs of states on the two occasions. The likelihood ratio test of item invariance can be formulated in terms of the expanded model. The constrained model representing the null hypothesis is that for any given state in the original model, parameter values for the same items in the two administrations are identical. The alternative general model allows the parameters to differ between the two administrations. However, it imposes the same constraints on parameters within each administration that were imposed in the original model. The number of parameters necessary to specify the distribution over states is the same for both models; hence, the degrees of freedom for the likelihood ratio chi-squared test is $K = 2K-K$, where K is the number of free item parameters in the original model.

Some unresolved problems. It would be nice if all the problems in developing a latent class model had straightforward solutions like the problems we have discussed so far. Unfortunately, this is not the case. There are a few issues that require a degree of subjective judgement on the part of the investigator. The lack of a general solution to these problems is not fatal to the process of model development, but it does mean that not much of value can be said about their solution out of the

context of the particular problem. It will be useful to at least describe some of these problems here, and discuss in the next section how they were treated in the signed-number context.

We need to deal with the following three problems:

1. How to compare two models which are not hierarchically related;
2. How to introduce an element of flexibility into the use of the likelihood ratio chi-squared statistics;
3. How to deal with the problem of relative, but nonglobal, maxima in the estimation algorithm.

The problem of comparison of two models which are not hierarchically related arises when we have to compare a constrained model with, say, five states, with an unconstrained model with four states. The mathematical rationale for the likelihood ratio test breaks down because neither model is a special case of the other.

The need for judgement in the use of hypothesis tests arises because the tests can be sensitive to real, but practically trivial, deviations from the null hypothesis. Blind application of the tests can lead one to employ a model with more parameters than one can accurately calibrate. It is often better to tolerate a little lack of fit in order to gain some accuracy in calibration. But what constitutes just a little lack of fit is a matter of judgement.

The first step in the parameter estimation algorithm is to specify suitable starting values for the parameters. This seemingly innocuous step can be the source of some difficulty. In a world designed for

solving estimation problems, the algorithm would eventually converge to same parameter values, no matter where it started. The algorithm given and most algorithms that would be practical alternatives to it are hill-climbing procedures. That is, they evaluate the behavior of the likelihood function at a given point; if the function is a local maximum, the algorithm terminates. If the function is not a local maximum at the point, then the algorithm in effect asks, "What is the best way up from here?". Thus, the algorithm eventually climbs to the top of whatever hill it happens to be on. If there is only one hill, the procedure works fine. If there are many hills, one must start on the biggest one, somehow. In latent class structure estimation problems, there definitely can be many hills, i.e., relative, but nonglobal, maxima. Thus, the choice of starting values can be crucial. This general problem can occur in a wide variety of situations, so it is an active area of current research in mathematical optimization theory, e.g. see Archetti and Szego (1980). Effective systematic procedures may eventually emerge for dealing with it. In the meantime, investigators must deal with it as best they can, trying out a reasonable variety of starting values to avoid getting stuck with a false maximum.

The development of models for signed-number addition

Determining the number of states. The results of chi-squared tests for adding states to unconstrained models for the two components of signed-number addition responses are given in Table 2. New states yield significant improvement up to four states for the absolute value component model and up to six states for the sign component model. Ideally, we should keep trying additional states until addition of another state

yields an insigificant chi-square. Unfortunately, the four-state model has $4 \times 20 + 3 = 83$ parameters, which nearly equals the number of distinct response patterns in the data from our sample of 129 subjects. The six-state model for the sign component has 125 parameters, which clearly exceeds the number of distinct response patterns observed. Thus, the model is overdetermined and we know, therefore, that the parameter estimates are not unique. There are other combinations of the conditional probabilities which would yield the same maximum value for the likelihood function. This makes interpretation of the parameters problematic, to say the least. The only point in continuing to add states to the respective models was to show that at least the number of states posited in the specific models described earlier in the paper would be required to account for the data. To obtain models whose parameter estimates are stable enough to be interpretable, it will be necessary to constrain the parameter values in some reasonable way.

INSERT TABLE 2 ABOUT HERE

Testing constrained models. Using four states for the absolute value component and six states for the sign component, we compare the performance of three models for each component: the unconstrained model, the model with the weak constraint that items within types are assumed to be equivalent, and the more constrained specific models summarized earlier in Table 1. The results of the chi-squared tests for these comparisons are presented in Table 3. We reject the null hypothesis that the more constrained model is adequate in every comparison.

Table 4 gives the estimated conditional proabilities of correct response, given state, for each item-state combination in the unconstrained and equivalent items models, for the absolute value component. In all but four of the item-type/state combinations, the single value for the equivalent items model falls in the range of the corresponding four parameter estimates for the unconstrained model. On the whole, the equivalent items estimates represent the unconstrained reasonably well, though one can see the discrepencies which lead to rejection of the equivalent items model by the chi-squared test. The instability of the unconstrained model parameter estimates due the small size of the sample is shown by the numerous zero's and one's among these estimates.

INSERT TABLES 3 AND 4 ABOUT HERE

Table 5 gives three sets of item parameter estimates for the absolute value component. Sets labelled C and C' are outputs of the maximum likelihood estimation algorithm. Set C contains the genuine maximum likelihood estimates for the constrained model in Table 1 and C' contains estimates which yield a local, but nonglobal, maximum to the likelihood function for the same model. Set E contains the maximum likelihood estimates for the more general model whose only constraint is that items within types must be equivalent.

Sets C and E are very similar, except for a couple of minor discrepencies in the parameters for the systematic states. The model E parameters for type "-L+-S" deviate from what is required by the

constraints of the specific model: they are a little too small in the "Mastery" and "Always add" states, and a little bit too large in the "Always subtract" state. The set E parameters in the other four types are relatively homogeneous for the "Mastery" and "Always subtract" states, but an interesting deviation occurs in the "Always add" state. Subjects in the latter state are more likely to deviate from their usual rule (and thereby respond correctly) on types "L+-S" and "-S+L" than on the types where the larger number has a negative sign. Note that "L+-S" and "-S+L" fit the pattern of ordinary subtraction problems with which the students are very familiar: it makes sense that their tendency to add would weaken on these items.

INSERT TABLE 5 ABOUT HERE

The parameter estimates in set C' arose out of experimentation with different starting values in the estimation algorithm to make sure that set C yielded a true maximum in the likelihood function. The original set of starting values, which yielded set C, were selected to fit the constraints of the model. All probabilities in the fourth state, the "guessing" state, were started at .50. In the systematic states, high values were started at .95 and low values at .05. Since the set C does not differ very dramatically from the starting values, it seemed plausible that we might have started out near a local maximum and been trapped. In order to avoid this possibility, a set of starting values which does not even conform to the constraints of the model was tried: all item parameters within each state were started at the same value. The common

values for the four respective states were .95, .80, .65, and .50. In retrospect, this was not a wise choice. According to the model, the state with lowest average parameter values is not the random state, but state 3, the "Always add" state. In that state, subjects make systematic errors on four of the five types of items, and therefore get fewer items correct than if they were guessing.

Examination of Table 5 reveals that sets C, C', and E are all fairly similar for the first two states, but an odd switch takes place in the third and fourth states. Both in terms of item parameters and state probabilities, the parameters in set C' in state 3 correspond to state 4 parameters in set C and vice versa. The C' parameters in state 4 look even more like the E parameters in state 3. In effect, the choice of starting values made C' good at matching the 11-16% of the subjects in the original state 3 at the expense of not matching as well the 36-38% in state 4. Thus, we wound up demonstrating how one could be led to a false maximum, but not quite the way we had in mind. It is interesting to note that it was not the starting values per se which led the algorithm to converge to a nonglobal maximum in this case, but the association of starting values with states. Using the same starting values, but assigning the value .50 to items in the third state and .65 to items in the fourth state leads the algorithm to converge to the parameters values in set C, the true maximum likelihood values.

Let us summarize the results obtained so far in developing the signed-number addition models. We need at least the number of states proposed in Table 1: four states for the absolute value component, six

states for the sign component. The unconstrained four-state model for the absolute value component fits the data significantly better than the equivalent items model and the equivalent items model fits better than the constrained model given in Table 1. The results of the goodness-of-fit tests notwithstanding, the parameters of the constrained model represent the unconstrained parameters quite well. An example was given to show that the problem of the estimation algorithm converging to a false maximum likelihood solution is not just an abstract possibility, but a real problem which users of the estimation algorithm need to keep in mind when selecting starting values.

All the points made here concerning the absolute value component apply as well to the sign component. There is nothing new to be learned by presenting the results in detail. However, it may be of interest to examine the match between observed and predicted values for the five subtests consisting of four equivalent items of each of the respective types.

Predictions concerning descriptive statistics. The predictions of any latent class model concerning the expected values, variances, and covariances of the responses can be expressed simply in matrix notation, which we now define. Let

P = states-by-items matrix of conditional probabilities of correct response to each item, given state;

$\underline{\nu}$ = column vector of state probabilities;

W = $\mathrm{diag}(\underline{\nu}) - \underline{\nu}\underline{\nu}'$, the covariance matrix of state membership, considered as a multinomial random variable;

D = $\mathrm{diag}(\sum_k v_k p_{kj}(1-p_{kj}))$, the diagonal matrix whose diagonal elements represent, for each item, the weighted average over states of the conditional variances of the responses to the items, given state;

$\underline{\mu} = E(\underline{x})$, the vector of expected values of the responses to the respective items;

Σ = interitem covariance matrix;

$\underline{c}$ = vector of coefficients used to define a linear combination of item responses, $\underline{c}'\underline{x} = \sum_j c_j x_j$.

The matrices P, W, and D, and the vector $\underline{v}$ are either given directly by or easily computed from the parameter estimates. Then the predicted vector of expected values of the item responses is given by

$$\underline{\mu} = P'\underline{v} , \tag{15}$$

and the predicted covariance matrix is

$$\Sigma = P'WP + D . \tag{16}$$

The predicted mean and variance of a linear combination of responses defined by a coefficient vector $\underline{c}$ are given by

$$E(\underline{c}'\underline{x}) = \underline{c}'\underline{\mu} \tag{17}$$

and

$$\mathrm{var}(\underline{c}'\underline{x}) = \underline{c}'\Sigma\underline{c}, \tag{18}$$

respectively. The predicted covariance between two linear combinations of responses defined by coefficient vectors $\underline{c}_1$ and $\underline{c}_2$ is

$$\text{cov}(\underline{c}_1'\underline{x},\underline{c}_2'\underline{x}) = \underline{c}_1'\Sigma\underline{c}_2. \tag{19}$$

Observed and predicted values of the means, standard deviations, and correlations between the numbers of errors on the respective item types are given in Table 6, for the two response components. The overall agreement with the predictions is quite good. It is particularly noteworthy that the constrained model is able to account for the pattern of positive and negative correlations in the data. The existence of negative correlations shows that the data cannot in principle be explained by a unidimensional item response theory model, according to a theorem of Holland (1981). It is hard to imagine a simpler alternative than the proposed latent class model which would be able to represent the data adequately.

INSERT TABLE 6 ABOUT HERE

There are two groups represented in this data which differ with respect to drills on signed-number addition which they received. The first group received "Postman Stories" and the second "Number Line Method". The two groups are very similar in their performance on the sign component, so the results are not given here for the separate groups on that component. The results are quite different for the two groups on the absolute value component, as can be seen in Table 7.

Testing the invariance of item parameters between two groups. The mean number of errors is smaller for the Postman Stories group than for the Number Line Method group on every type of item. The standard deviations are similar in the two groups on all types, but there are noteworthy differences between the groups in the correlations between types. The correlations are all positive in the Postman Stories group. The correlations are uniformly lower in the Number Line group - markedly so in the case of correlations between the "-L+-S" item type and the other types, where they are negative. Even though the correlations in the two groups differ substantially, both sets are in good agreement with the predictions of the model. The two-sample test for invariance of parameters described above yields a chi-squared value of 17.90, which is not significant at the .05 level with 11 degrees of freedom.

INSERT TABLE 7 ABOUT HERE

Evidently, the model can explain the differences between the two groups entirely in terms of differences in distribution of subjects over states. The estimated distributions over the Mastery, Always Add, Always Subtract, and Random states are (.34, .07, .24, .35) in the Postman Stories group and (.08, .11, .24, .57) in the Number Line Method group. Thus, the two groups have similar proportions in the Always Add and Always Subtract states, but they differ in their proportions in the other two states. The Postman Stories group has an even split of the remaining subjects between the Mastery and Random states, whereas the Number Line group has the large majority of its remaining subjects in the Random

group. This clearly accounts for the smaller number of errors in the Postman Stories group. To see how it accounts for the differences between the groups in the correlations between the "-L+-S" type and the other types, consider the composition of a typical 2×2 table for computing the phi-coefficient between an "-L+-S" item and one of the other items.

The "-L+-S" item type is the only one on the test which calls for adding the absolute values to determine the magnitude of the answer. Subjects in the Always Add or Always Subtract states therefore tend to get the absolute value of the "-L+-S" item right and the other item wrong, or vice versa. Since the two groups have roughly the same proportions of their subjects in these two states, they both have significant frequencies in the off-diagonal cells of the 2×2 table coming from these states. In the Postman Stories group, these off-diagonal frequencies are offset by a substantial frequency in the cell for correct responses to both items contributed by subjects in the Mastery state. However, this contribution is lacking in the Number Line group. The result is the negative correlations observed in the Number Line group and the small positive correlations observed in the Postman Stories group.

Implications of the signed-number addition results. The signed-number addition models demonstrate that one can develop models in the latent class structure framework which possess two important properties. First, these models can provide a simple, meaningful representation of subject state capable of dealing with test data which is not unidimensional. Second, this representation may also be capable of representing differences between groups or changes within groups over time, even though the variance-covariance structure of the responses

differs in the two groups or at the different times. It is beyond the scope of this paper to deal in any detail with applications of latent class models in test development and interpretation. However, a few significant potential applications might at least be mentioned.

When test data are not unidimensional, but a meaningful latent class model can be developed to represent the data, it would be natural to develop diagnostic assessment procedures to replace or supplement the usual numerical scoring of the test. If the test results are intended to be used to make mastery/nonmastery decisions, the model can be used to develop better criteria and bases for setting criteria.

One problem that plagues the interpretation of changes in achievement test scores is that the factor structure or dimensionality of the test changes when the average level of achievement of the group being tested changes (Gialluca and Weiss, 1981; Tatsuoka and Birenbaum, 1979). While the data just described comes from two groups tested at the same time, rather than one group measured at two different times, the major difference between the groups is what we would expect if the item parameters stayed the same over two testings of a single group, but the distribution over states changed. This suggests that it will be possible to interpret changes between testings in terms of the model presented here, even though the factor structure of the test changes between testings. Development of model-based procedures for assessing change would be a natural research direction to follow.

The test for parameter invariance between two or more groups has many potential applications in the study of group differences, such as studies of test bias and cross-cultural psychology.

SUMMARY

This paper has been concerned with the use of latent class structure as a modelling framework for tests in which much of the data conforms to a relatively small number of systematic patterns. Application of this framework to the analysis of tests has been limited because available parameter estimation algorithms can only handle a relatively small number of items. The paper has proposed a modification of the constrained maximum likelihood estimation algorithm of Goodman (1974) which should extend the applicability of that algorithm to tests with many more items. A number of special cases of the latent class structure model were reviewed, including models which we propose for signed-number addition test responses. A series of problems which arise in the development of any latent class model was discussed and likelihood ratio tests for dealing with the problems were presented. A few problems with no satisfactory general solution were noted. The proposed procedures were used to develop and test models for signed-number addition test data gathered by Tatsuoka and Birenbaum (1979). The models were shown to represent most features of the data well, including some significantly negative interitem correlations and differences between groups in covariance structure. Negative interitem correlations imply that the data are not unidimensional. The ability of these simple models to represent non-unidimensional data, and to represent differences between groups in covariance structure and means simultaneously, suggests that the latent class structure framework may be very useful in dealing with a number of troublesome psychometric issues.

TABLE 1. Models specifying conditional probabilities of correct response for the two components of signed-number addition problems.*

(a) Sign Component

State of Subject	Item Type L+-S	-L+S	-L+-S	-S+L	S+-L
1. Correct rule	$1-\varepsilon_1$	$1-\varepsilon_1$	$1-\varepsilon_1$	$1-\varepsilon_1$	$1-\varepsilon_1$
2. Always positive	$1-\varepsilon_2$	ε_2	ε_2	$1-\varepsilon_2$	ε_2
3. Always negative	ε_3	$1-\varepsilon_3$	$1-\varepsilon_3$	ε_3	$1-\varepsilon_3$
4. Sign of first	$1-\varepsilon_4$	$1-\varepsilon_4$	$1-\varepsilon_4$	ε_4	ε_4
5. Sign of second	ε_5	ε_5	$1-\varepsilon_5$	$1-\varepsilon_5$	ε_5
6. Random guessing	p_{r1}	p_{r2}	p_{r3}	p_{r4}	p_{r5}

(b) Absolute Value Component

State of Subject	Item Type L+-S	-L+S	-L+-S	-S+L	S+-L
1. Correct rule	$1-\varepsilon_1$	$1-\varepsilon_1$	$1-\varepsilon_1$	$1-\varepsilon_1$	$1-\varepsilon_1$
2. Always adds	ε_2	ε_2	$1-\varepsilon_2$	ε_2	ε_2
3. Always subtracts	$1-\varepsilon_3$	$1-\varepsilon_3$	ε_3	$1-\varepsilon_3$	$1-\varepsilon_3$
4. Random guessing	p_{r1}	p_{r2}	p_{r3}	p_{r4}	p_{r5}

*The same generic symbols are used for both components--ε_i's for errors in execution of a systematic rule, P_{rj}'s for probabilities of correct random guesses. It should be understood that ε_i's and P_{rj}'s for the sign component are distinct from ε_i's and P_{rj}'s for the absolute value component.

TABLE 2. Likelihood ratio chi-squares for determining the number of states for the two components of signed-number addition problems, using unconstrained models.

(a) Sign Component

Number of States	-Log L	χ^2	df	p
2	1334.83			
3	1140.51	388.64	21	<.001
4	1101.11	78.80	21	<.001
5	1060.08	82.06	21	<.001
6	1038.75	42.66	21	<.005

(b) Absolute Value Component

Number of States	-Log L	χ^2	df	p
2	1370.42			
3	1306.28	128.28	21	<.001
4	1242.22	128.12	21	<.001

Table 3. Likelihood ratio chi-squares for a hierarchy of constrained models for the sign and absolute value components

		Model	-logL	χ^2	df	p
Sign Component	1.	Constrained, six state	1131.90			
	2.	Equivalent items	1101.62	60.56	20	<.001
	4.	Unconstrained six state	1042.33	118.58	90	<.05
Absolute Value Component	1.	Constrained, four state	1320.47			
	2.	Equivalent items	1294.56	49.82	12	<.001
	3.	Unconstrained,	1242.22	104.64	60	<.001

Table 4. Comparison of parameters in the absolute value component model with equivalent items constraint, Model E, with unconstrained model parameters, Model U.

	State 1		State 2		State 3		State 4	
Item	E	U	E	U	E	U	E	U
1. (L+-S)	.98	1.00	.87	.81	.20	.39	.64	.74
2.	.98	.91	.87	.91	.20	.14	.64	.75
3.	.98	.85	.87	.89	.20	.09	.64	.53
4.	.98	.97	.87	1.00	.20	.25	.64	.77
5. (-L+S)	.98	.97	.90	.86	.03	.00	.39	.42
6.	.98	.97	.90	.94	.03	.00	.39	.43
7.	.98	1.00	.90	.86	.03	.23	.39	.46
8.	.98	.91	.90	1.00	.03	.06	.39	.45
9. (-L+-S)	.86	.60	.19	.17	.79	.57	.35	.37
10.	.86	.83	.19	.13	.79	.76	.35	.36
11.	.86	.88	.19	.00	.79	.57	.35	.37
12.	.86	.81	.19	.11	.79	.66	.35	.39
13. (-S+L)	.98	1.00	.92	.97	.21	.25	.40	.63
14.	.98	1.00	.92	.97	.21	.25	.40	.50
15.	.98	.93	.92	.94	.21	.11	.40	.29
16.	.98	1.00	.92	.81	.21	.09	.40	.43
17. (S+-L)	.95	1.00	.83	.70	.02	.14	.51	.71
18.	.95	.88	.83	.93	.02	.00	.51	.53
19.	.95	.82	.83	.97	.02	.19	.51	.44
20.	.95	.91	.83	.87	.02	.00	.51	.44
State Probability	.20	.25	.32	.24	.12	.16	.36	.34

Table 5. Three sets of estimates for the parameters of models for the absolute value component: set C consists of genuine maximum likelihood estimates for the constrained model; set C' consists of another set of solutions to the maximum likelihood equations which yields a nonglobal maximum for the likelihood function; set E contains the maximum likelihood estimates for the equivalent items model.

		Item Type					
State		L+-S	-L+S	-L+-S	-S+L	S+-L	State Probability
1.	C	.94	.94	.94	.94	.94	.21
	C'	.95	.95	.95	.95	.95	.20
	E	.98	.98	.86	.98	.95	.20
2.	C	.88	.88	.12	.88	.88	.30
	C'	.89	.89	.11	.89	.89	.28
	E	.87	.90	.19	.92	.83	.32
3.	C	.12	.12	.88	.12	.12	.11
	C'	.56	.56	.44	.56	.56	.36
	E	.20	.03	.79	.21	.02	.12
4.	C	.63	.39	.36	.40	.50	.38
	C'	.22	.07	.64	.18	.08	.16
	E	.64	.39	.35	.40	.51	.36

TABLE 6. Observed and predicted values (in parentheses) of means, standard deviations, and correlations between the errors on the respective item types (N=129).

(a) Sign Component

	L + -S	-S + L	-L + -S	S + -L	-L + S
L + -S	1.00	-.12(-.04)	-.20(-.11)	.55(.61)	-.14(-.19)
-S + L		1.00	.72(.54)	-.24(-.15)	.57(.55)
-L + -S			1.00	-.22(-.14)	.66(.57)
S + -L				1.00	.03(.05)
-L + S					1.00
m_j	1.08(---)	1.19(---)	.74(---)	1.24(---)	1.12(---)
s_j	1.36(1.41)	1.45(1.40)	1.17(1.18)	1.41(1.45)	1.40(1.36)

(b) Absolute Value Component

	L + -S	-S + L	-L +-S	S + -L	-L + S
L + -S	1.00	.61(.57)	-.07(-.10)	.61(.58)	.73(.57)
-S + L		1.00	.03(.00)	.78(.64)	.62(.65)
-L + -S			1.00	-.02(-.03)	-.13(-.02)
S + -L				1.00	.59(.63)
-L + S					1.00
m_j	1.08(---)	1.49(---)	2.19(---)	1.36(---)	1.43(---)
s_j	1.28(1.18)	1.56(1.36)	1.44(1.48)	1.44(1.30)	1.46(1.34)

TABLE 7. Summary statistics (with predicted values in parentheses) on the absolute value component, for the two separate treatment groups.

Group 1--Postman Stories (N = 67)

	L + -S	-S + L	-L + -S	S + -L	-L + S
L + -S	1.00	.64(.59)	.07(.12)	.74(.58)	.78(.59)
-S + L		1.00	.19(.29)	.80(.66)	.63(.71)
-L + -S			1.00	.15(.21)	.04(.27)
S + -L				1.00	.59(.65)
-L + S					1.00
m_j	.88(---)	1.27(---)	2.03(---)	1.06(---)	1.19(---)
s_j	1.11(1.11)	1.23(1.39)	1.24(1.64)	1.18(1.23)	1.18(1.33)

Group 2--Number Line Stories (N = 62)

	L + -S	-S + L	-L +-S	S + -L	-L + S
L + -S	1.00	.57(.54)	-.28(-.33)	.46(.54)	.68(.54)
-S + L		1.00	-.20(-.27)	.75(.59)	.58(.59)
-L + -S			1.00	-.28(-.27)	-.38(-.27)
S + -L				1.00	.55(.59)
-L + S					1.00
m_j	1.31(---)	1.73(---)	2.35(---)	1.69(---)	1.69(---)
s_j	1.15(1.23)	1.26(1.33)	1.14(1.35)	1.19(1.32)	1.22(1.32)

Reference Notes

Tatsuoka, K., and Birenbaum, M. The danger of relying solely on diagnostic adaptive testing when prior and subsequent instructional methods are different. Computer-based Education Research Laboratory Report E-S. University of Illinois, 1979.

References

Archetti, F. and Szego, G.P. Global optimization algorithms. Lecture 17 in Dixon, L.C.W., Spedicato, E., ad Szego, G.P. Nonlinear optimization, theory and algorithms. Birkhauser, Boston, 1980.

Bock, R.D., and Aitkin, M. Marginal maximum likelihood estimation of item parameters: applciation of an EM algorithm. Psychometrika, 1981, 46, 443-459.

Bock, R.D., and Lieberman, M. Fitting a response model for n dichotomously scored items. Psychometrika, 1970, 35, 179-197.

Brown, J.S., and Burton, R.R. Diagnostic models for procedural bugs in basic mathematical skills. Cognitive Science, 1978, 2, 155-192.

Clogg, C.C., and Sawyer, D.O. A comparison of alternative models for analyzing the scalability of response patterns. In S. Leinhardt (Ed.), Sociological Methodology 1981. San Francisco: Jossey-Bass, 1981.

Dayton, C.M., and Macready, G.B. A probabilistic model for validation of behavioral hierarchies. Psychometrika, 1976, 41, 189-204.

Dayton, C.M., and Macready, G.B. A scaling model with response errors and intrinsically unscalable respondents, Psychometrika, 1980, 45, 343-356.

Dempster, A.P., Laird, N.M., and Rubin, D.B. Maximum likelihood from incomplete data via the EM algorithm (with Discussion). Journal of the Royal Statistical Society, Series B, 1977, 39, 1-38.

Goodman, L.A. The analysis of qualitative variables when some of the variables are unobservable. Part I--A modified latent structure approach. American Journal of Sociology, 1974, 79, 1179-1259.

Goodman, L.A. A new model for scaling response patterns: An application of the quasi-independence concept. Journal of the American Statistical Association, 1975, 70, 755-768.

Holland, P.W. When are item response models consistent with observed data? Psychometrika, 1981, 46, 79-91.

Lazarsfeld, P.F., and Henry, N.W. Latent structure analysis. Boston: Houghton-Mifflin, 1968.

McHugh, R.B. Efficient estimation and local identification in latent class analysis. Psychometrika, 1956, 21, 331-347.

Personnel Analysis Division,
AF/MPXA
5C360, The Pentagon
Washington, DC 20330

Air Force Human Resources Lab
AFHRL/MPD
Brooks AFB, TX 78235

Dr. Earl A. Alluisi
HQ, AFHRL (AFSC)
Brooks AFB, TX 78235

Dr. Erling B. Andersen
Department of Statistics
Studiestraede 6
1455 Copenhagen
DENMARK

Dr. Phipps Arabie
University of Illinois
Department of Psychology
603 E. Daniel St.
Champaign, IL 61820

Technical Director, ARI
5001 Eisenhower Avenue
Alexandria, VA 22333

Dr. Eva L. Baker
UCLA Center for the Study
of Evaluation
145 Moore Hall
University of California
Los Angeles, CA 90024

Dr. Isaac Bejar
Educational Testing Service
Princeton, NJ 08450

Dr. Menucha Birenbaum
School of Education
Tel Aviv University
Tel Aviv, Ramat Aviv 69978
ISRAEL

Dr. Arthur S. Blaiwes
Code N711
Naval Training Equipment Center
Orlando, FL 32813

Dr. R. Darrell Bock
University of Chicago
Department of Education
Chicago, IL 60637

Cdt. Arnold Bohrer
Sectie Psychologisch Onderzoek
Rekruterings-En Selectiecentrum
Kwartier Koningen Astrid
Bruijnstraat
1120 Brussels, BELGIUM

Dr. Robert Breaux
Code N-095R
NAVTRAEQUIPCEN
Orlando, FL 32813

Dr. Robert Brennan
American College Testing
Programs
P. O. Box 168
Iowa City, IA 52243

Dr. Patricia A. Butler
NIE Mail Stop 1806
1200 19th St., NW
Washington, DC 20208

Mr. James W. Carey
Commandant (G-PTE)
U.S. Coast Guard
2100 Second Street, S.W.
Washington, DC 20593

Dr. James Carlson
American College Testing
Program
P.O. Box 168
Iowa City, IA 52243

Dr. John B. Carroll
409 Elliott Rd.
Chapel Hill, NC 27514

Dr. Robert Carroll
NAVOP 01B7
Washington, DC 20370

Dr. Norman Cliff
Department of Psychology
Univ. of So. California
University Park
Los Angeles, CA 90007

Director,
Manpower Support and
Readiness Program
Center for Naval Analysis
2000 North Beauregard Street
Alexandria, VA 22311

Dr. Stanley Collyer
Office of Naval Technology
Code 222
800 N. Quincy Street
Arlington, VA 22217-5000

Dr. Hans Crombag
University of Leyden
Education Research Center
Boerhaavelaan 2
2334 EN Leyden
The NETHERLANDS

CTB/McGraw-Hill Library
2500 Garden Road
Monterey, CA 93940

Dr. Dattprasad Divgi
Center for Naval Analysis
4401 Ford Avenue
P.O. Box 16268
Alexandria, VA 22302-0268

Dr. Hei-Ki Dong
Ball Foundation
800 Roosevelt Road
Building C, Suite 206
Glen Ellyn, IL 60137

Defense Technical
Information Center
Cameron Station, Bldg 5
Alexandria, VA 22314
Attn: TC
(12 Copies)

Dr. Stephen Dunbar
Lindquist Center
for Measurement
University of Iowa
Iowa City, IA 52242

Dr. James A. Earles
Air Force Human Resources Lab
Brooks AFB, TX 78235

Dr. Kent Eaton
Army Research Institute
5001 Eisenhower Avenue
Alexandria, VA 22333

Dr. John M. Eddins
University of Illinois
252 Engineering Research
Laboratory
103 South Mathews Street
Urbana, IL 61801

Dr. Susan Embretson
University of Kansas
Psychology Department
Lawrence, KS 66045

ERIC Facility-Acquisitions
4833 Rugby Avenue
Bethesda, MD 20014

Dr. Benjamin A. Fairbank
Performance Metrics, Inc.
5825 Callaghan
Suite 225
San Antonio, TX 78228

Dr. Leonard Feldt
Lindquist Center
for Measurement
University of Iowa
Iowa City, IA 52242

Dr. Richard L. Ferguson
American College Testing
Program
P.O. Box 168
Iowa City, IA 52240

Dr. Gerhard Fischer
Liebiggasse 5/3
A 1010 Vienna
AUSTRIA

Prof. Donald Fitzgerald
University of New England
Department of Psychology
Armidale, New South Wales 2351
AUSTRALIA

Mr. Paul Foley
Navy Personnel R&D Center
San Diego, CA 92152

Dr. Carl H. Frederiksen
McGill University
3700 McTavish Street
Montreal, Quebec H3A 1Y2
CANADA

Dr. Robert D. Gibbons
University of Illinois-Chicago
P.O. Box 6998
Chicago, IL 69680

Dr. Janice Gifford
University of Massachusetts
School of Education
Amherst, MA 01003

Dr. Robert Glaser
Learning Research
& Development Center
University of Pittsburgh
3939 O'Hara Street
Pittsburgh, PA 15260

Dr. Bert Green
Johns Hopkins University
Department of Psychology
Charles & 34th Street
Baltimore, MD 21218

Dr. Ronald K. Hambleton
Prof. of Education & Psychology
University of Massachusetts
at Amherst
Hills House
Amherst, MA 01003

Ms. Rebecca Hetter
Navy Personnel R&D Center
Code 62
San Diego, CA 92152

Dr. Paul W. Holland
Educational Testing Service
Rosedale Road
Princeton, NJ 08541

Prof. Lutz F. Hornke
Universitat Dusseldorf
Erziehungswissenschaftliches
Universitatsstr. 1
Dusseldorf 1
WEST GERMANY

Dr. Paul Horst
677 G Street, #184
Chula Vista, CA 90010

Mr. Dick Hoshaw
NAVOP-135
Arlington Annex
Room 2834
Washington, DC 20350

Dr. Lloyd Humphreys
University of Illinois
Department of Psychology
603 East Daniel Street
Champaign, IL 61820

Dr. Steven Hunka
Department of Education
University of Alberta
Edmonton, Alberta
CANADA

Dr. Huynh Huynh
College of Education
Univ. of South Carolina
Columbia, SC 29208

Dr. Robert Jannarone
Department of Psychology
University of South Carolina
Columbia, SC 29208

Dr. Douglas H. Jones
Advanced Statistical
Technologies Corporation
10 Trafalgar Court
Lawrenceville, NJ 08148

Dr. G. Gage Kingsbury
Portland Public Schools
Research and Evaluation Department
501 North Dixon Street
P. O. Box 3107
Portland, OR 97209-3107

Dr. William Koch
University of Texas-Austin
Measurement and Evaluation
Center
Austin, TX 78703

Dr. H. Wallace Sinaiko
Manpower Research
and Advisory Services
Smithsonian Institution
801 North Pitt Street
Alexandria, VA 22314

Dr. Richard Sorensen
Navy Personnel R&D Center
San Diego, CA 92152

Dr. Paul Speckman
University of Missouri
Department of Statistics
Columbia, MO 65201

Dr. Martha Stocking
Educational Testing Service
Princeton, NJ 08541

Dr. Peter Stoloff
Center for Naval Analysis
200 North Beauregard Street
Alexandria, VA 22311

Dr. William Stout
University of Illinois
Department of Mathematics
Urbana, IL 61801

Maj. Bill Strickland
AF/MPXOA
4E168 Pentagon
Washington, DC 20330

Dr. Hariharan Swaminathan
Laboratory of Psychometric and
Evaluation Research
School of Education
University of Massachusetts
Amherst, MA 01003

Mr. Brad Sympson
Navy Personnel R&D Center
San Diego, CA 92152

Dr. Kikumi Tatsuoka
CERL
252 Engineering Research
Laboratory
Urbana, IL 61801

Dr. Maurice Tatsuoka
220 Education Bldg
1310 S. Sixth St.
Champaign, IL 61820

Dr. David Thissen
Department of Psychology
University of Kansas
Lawrence, KS 66044

Mr. Gary Thomasson
University of Illinois
Educational Psychology
Champaign, IL 61820

Dr. Robert Tsutakawa
The Fred Hutchinson
Cancer Research Center
Division of Public Health Sci.
1124 Columbia Street
Seattle, WA 98104

Dr. Ledyard Tucker
University of Illinois
Department of Psychology
603 E. Daniel Street
Champaign, IL 61820

Dr. Vern W. Urry
Personnel R&D Center
Office of Personnel Management
1900 E. Street, NW
Washington, DC 20415

Dr. David Vale
Assessment Systems Corp.
2233 University Avenue
Suite 310
St. Paul, MN 55114

Dr. Frank Vicino
Navy Personnel R&D Center
San Diego, CA 92152

Dr. Howard Wainer
Division of Psychological Studies
Educational Testing Service
Princeton, NJ 08541

Dr. Ming-Mei Wang
Lindquist Center
for Measurement
University of Iowa
Iowa City, IA 52242

Mr. Thomas A. Warm
Coast Guard Institute
P. O. Substation 18
Oklahoma City, OK 73169

Dr. Brian Waters
Program Manager
Manpower Analysis Program
HumRRO
1100 S. Washington St.
Alexandria, VA 22314

Dr. David J. Weiss
N660 Elliott Hall
University of Minnesota
75 E. River Road
Minneapolis, MN 55455

Dr. Ronald A. Weitzman
NPS, Code 54Wz
Monterey, CA 92152

Major John Welsh
AFHRL/MOAN
Brooks AFB, TX 78223

Dr. Rand R. Wilcox
University of Southern
California
Department of Psychology
Los Angeles, CA 90007

German Military Representative
ATTN: Wolfgang Wildegrube
Streitkraefteamt
D-5300 Bonn 2
4000 Brandywine Street, NW
Washington, DC 20016

Dr. Bruce Williams
Department of Educational
Psychology
University of Illinois
Urbana, IL 61801

Dr. Hilda Wing
Army Research Institute
5001 Eisenhower Ave.
Alexandria, VA 22333

Dr. Martin F. Wiskoff
Navy Personnel R & D Center
San Diego, CA 92152

Mr. John H. Wolfe
Navy Personnel R&D Center
San Diego, CA 92152

Dr. George Wong
Biostatistics Laboratory
Memorial Sloan-Kettering
Cancer Center
1275 York Avenue
New York, NY 10021

Dr. Wendy Yen
CTB/McGraw Hill
Del Monte Research Park
Monterey, CA 93940

Dr. Leonard Kroeker
Navy Personnel R&D Center
San Diego, CA 92152

Dr. Michael Levine
Educational Psychology
210 Education Bldg.
University of Illinois
Champaign, IL 61801

Dr. Charles Lewis
Faculteit Sociale Wetenschappen
Rijksuniversiteit Groningen
Oude Boteringestraat 23
9712GC Groningen
The NETHERLANDS

Dr. Robert Linn
College of Education
University of Illinois
Urbana, IL 61801

Dr. Robert Lockman
Center for Naval Analysis
4401 Ford Avenue
P.O. Box 16268
Alexandria, VA 22302-0268

Dr. Frederic M. Lord
Educational Testing Service
Princeton, NJ 08541

Dr. James Lumsden
Department of Psychology
University of Western Australia
Nedlands W.A. 6009
AUSTRALIA

Dr. William L. Maloy
Chief of Naval Education
and Training
Naval Air Station
Pensacola, FL 32508

Dr. Gary Marco
Stop 31-E
Educational Testing Service
Princeton, NJ 08451

Dr. Clessen Martin
Army Research Institute
5001 Eisenhower Blvd.
Alexandria, VA 22333

Dr. James McBride
Psychological Corporation
c/o Harcourt, Brace,
Javanovich Inc.
1250 West 6th Street
San Diego, CA 92101

Dr. Clarence McCormick
HQ, MEPCOM
MEPCT-P
2500 Green Bay Road
North Chicago, IL 60064

Mr. Robert McKinley
University of Toledo
Department of Educational Psychology
Toledo, OH 43606

Dr. Barbara Means
Human Resources
Research Organization
1100 South Washington
Alexandria, VA 22314

Dr. Robert Mislevy
Educational Testing Service
Princeton, NJ 08541

Headquarters, Marine Corps
Code MPI-20
Washington, DC 20380

Dr. W. Alan Nicewander
University of Oklahoma
Department of Psychology
Oklahoma City, OK 73069

Dr. William E. Nordbrock
FMC-ADCO Box 25
APO, NY 09710

Dr. Melvin R. Novick
356 Lindquist Center
for Measurement
University of Iowa
Iowa City, IA 52242

Director, Manpower and Personnel
Laboratory,
NPRDC (Code 06)
San Diego, CA 92152

Library, NPRDC
Code P201L
San Diego, CA 92152

Commanding Officer,
Naval Research Laboratory
Code 2627
Washington, DC 20390

Dr. James Olson
WICAT, Inc.
1875 South State Street
Orem, UT 84057

Office of Naval Research,
Code 1142PT
800 N. Quincy Street
Arlington, VA 22217-5000
(6 Copies)

Special Assistant for Marine
Corps Matters,
ONR Code 00MC
800 N. Quincy St.
Arlington, VA 22217-5000

Dr. Judith Orasanu
Army Research Institute
5001 Eisenhower Avenue
Alexandria, VA 22333

Wayne M. Patience
American Council on Education
GED Testing Service, Suite 20
One Dupont Circle, NW
Washington, DC 20036

Dr. James Paulson
Department of Psychology
Portland State University
P.O. Box 751
Portland, OR 97207

Dr. Roger Pennell
Air Force Human Resources
Laboratory
Lowry AFB, CO 80230

Dr. Mark D. Reckase
ACT
P. O. Box 168
Iowa City, IA 52243

Dr. Malcolm Ree
AFHRL/MP
Brooks AFB, TX 78235

Dr. Carl Ross
CNET-PDCD
Building 90
Great Lakes NTC, IL 60088

Dr. J. Ryan
Department of Education
University of South Carolina
Columbia, SC 29208

Dr. Fumiko Samejima
Department of Psychology
University of Tennessee
Knoxville, TN 37916

Mr. Drew Sands
NPRDC Code 62
San Diego, CA 92152

Dr. Robert Sasmor
Army Research Institute
5001 Eisenhower Avenue
Alexandria, VA 22333

Dr. Mary Schratz
Navy Personnel R&D Center
San Diego, CA 92152

Dr. W. Steve Sellman
OASD(MRA&L)
2B269 The Pentagon
Washington, DC 20301

Dr. Kazuo Shigemasu
7-9-24 Kugenuma-Kaigan
Fujusawa 251
JAPAN

Dr. William Sims
Center for Naval Analysis
4401 Ford Avenue
P.O. Box 16268
Alexandria, VA 22302-0268